Crowding Valuation in Urban Tram and Bus Transportation based on Smart Card Data

Yap, Menno; Cats, Oded; Yu, S; van Arem, Bart

Publication date
2017

Document Version
Accepted author manuscript

Published in
Thredbo 15

Crowding valuation in urban tram and bus transportation based on smart card data

Menno Yap[a*], Oded Cats[a], Shuixian Yu[a] and Bart van Arem[a]

[a] *Delft University of Technology, P.O. Box 5048, 2628 CN Delft, the Netherlands*

* Corresponding author: Delft University of Technology, department of Transport and Planning, Stevinweg 1, 2628 CN Delft, the Netherlands, +31 (0)15 27 89802, *M.D.Yap@TUDelft.nl*

Keywords:
Big Data; Crowding; Public Transport; Revealed Preference; Smart Card Data;

Classification codes (minimum of 3, maximum 6):
C51, C55, C81, R41

ABSTRACT

Crowding in public transport can be of major influence on passengers' travel experience and therefore affect route and mode choice. Therefore, it is important to understand how crowding in urban public transport is perceived by passengers. The availability of individual smart card transactions allows us to gain insights in revealed trade-offs between travel time, transfers, waiting time and crowding in public transport route choice, instead of using stated preference experiments. Our study is the first in which crowding valuation for urban tram and bus travelling is determined fully based on revealed preference data. Based on the estimated discrete choice model, we conclude that crowding plays a significant role in passengers' route choice in public transport. The crowding multiplier of in-vehicle time equals 1.16 when all seats are occupied. The crowding multiplier increases further from 1.16 to 1.40, in case of an average density of 4 standing passengers per m^2. Besides, our estimation results confirm the existence of a tram bonus, in which in-vehicle time in a tram is perceived 0.6 times the in-vehicle time perceived in a bus. The insights gained from our study can support the decision-making process, by quantifying benefits of measures aiming to reduce crowding levels.

1. Introduction

Crowding in public transport can be of major influence on passengers' travel experience and therefore affect route and mode choice. Because of the expected increasing concentration of activities within urban agglomerations in many countries worldwide, crowding is expected to become an even more dominant factor in urban public transportation in the future. Therefore, it is important to understand how crowding in urban public transport is perceived by passengers. This facilitates policy makers in the decision-making process regarding topics as modal split, urban accessibility, liveability and sustainability. It can also contribute to the quantification of societal benefits of measures aiming to reduce crowding levels (see for example Prud'Homme et al. 2012; Haywood & Koning 2015; Cats et al. 2016).

Over the years, there are several studies performed aiming to better understand the valuation of crowding in public transport. Extensive overviews and meta analyses regarding crowding valuation studies can for example be found in Wardman & Whelan (2011) and in Li & Hensher (2011). The majority of the studies described in these papers use stated preference (SP) experiments to estimate crowding valuation. In some studies SP results are validated against a usually limited set of revealed preference (RP) data, for example by using surveys, passenger counts or cameras (e.g. Kroes et al. 2014). These SP based results are applied in public transport models aiming to improve passenger assignment and predictions, see for example Hamdouch et al. (2011), Schmöcker et al. (2011), Nuzzolo et al. (2012), Pel et al. (2014), Cats et al. (2016) and Van Oort et al. (2016). It is however known that there can be a discrepancy between stated choices of respondents in SP experiments, compared to revealed choice behaviour in reality. SP results can therefore be problematic in terms of realism. Until recently, the use of RP data could be troublesome especially regarding data availability and variability. However, the availability of automatic fare collection (AFC), automatic vehicle location (AVL) and automatic passenger count (APC) systems in public transport the last years enables the estimation of crowding valuation based on large scale revealed preference data. The availability of individual smart card transactions allows us to gain insights in revealed trade-offs between (for example) travel time, transfers, waiting time and crowding in public transport route choice.

Only recently, Hörcher et al. (2017) estimated discrete choice models incorporating crowding valuation in metro systems fully based on revealed preference data, by integrating AFC and AVL data. In this study, 32 origin-destination (OD) pairs of the MTR metro network of Hong Kong are used to estimate the valuation of the standing probability and crowding density (expressed in the number of standing passengers per square meter). Since the AFC system in Hong Kong is a closed, station-based system where passengers have to tap-in and tap-out at the metro station, the exact route and trip choice needs to be inferred using a passenger-to-train assignment method. In our study, we elaborate on the work by Hörcher et al. (2017), by estimating crowding valuation during urban tram and bus journeys. To the end, we use AFC and AVL data of the total urban public transport network of The Hague, the Netherlands. Since passengers are required to tap-in and tap-out within each tram or bus vehicle, no inference of the exact route or the exact vehicle choice is required. By merging AFC and AVL data, it is possible to directly determine the exact route and vehicle each passenger used, and to determine the stop-to-stop vehicle occupancy for each vehicle trip separately. This allows us to use to direct smart card data to estimate crowding valuation,

without the need to apply a relatively complex assignment procedure. Since the urban tram and bus network has a higher density compared to a metro network, it is possible to incorporate a larger number of OD pairs in our data set of which route choice might partly be explained by expected crowding levels on the different route alternatives. Besides, next to the crowding estimation based on the density of standing passengers, we explicitly incorporate the estimation of a crowding in-vehicle time multiplier for occupancies lower than the seat capacity as well.

This paper is structured as follows. Chapter 2 discusses the methodology applied regarding data processing (2.1), transfer inference (2.2), selection of OD pairs (2.3), determination of attributes and attribute levels (2.4) and model formulation (2.5). Chapter 3 shows the estimation results and discusses implications of these results. In chapter 4, conclusions and recommendations for further research are formulated.

2. Methodology

2.1 Data processing

We use the urban public transport network of The Hague, the Netherlands, to estimate public transport crowding valuation. This network consists of 12 tram lines and 8 bus lines, operated by the urban public transport operator HTM. Two of these 12 tram lines function as light rail service on the agglomerative network level, connecting The Hague with the satellite city of Zoetermeer. The other 10 tram lines and all bus lines function as urban lines within The Hague. In The Hague, there is no metro system available. When travelling by light rail, tram or bus in the Netherlands, passengers are required to tap-in and tap-out at devices located within each vehicle. This means that there is a closed, entry-exit fare system applied in The Hague. This is different from most urban public transport systems in the world, in which especially for busses often an open, entry-only system with flat fare structure is applied. This can for example be seen in London (Gordon et al. 2013) or in Santiago, Chile (Munigaza & Palma 2012). The fare system as applied in the Netherlands means that for each journey leg made by each individual the boarding time and boarding location, as well as the alighting time and location, are known. Besides, for each smart card transaction the line number, vehicle number, trip number and smart card number are known. This means that each journey leg with its corresponding transaction information appears as a separate row in the AFC dataset. During an average working day, more than 300.000 AFC transactions are made on the urban public transport network in The Hague. In the Netherlands, the AFC data is closed data owned by the public transport operator. The AVL data on the other hand is open data and publicly available. The AVL dataset contains the scheduled and realized arrival time and departure time of each vehicle trip at each station.

The original dataset used for this study contains all AFC and AVL data of 28 days from November 2 – November 29, 2015. This means that about 7.4 million AFC transactions and about 3.1 million AVL registrations are available in the dataset. In the data processing phase, we applied the following steps.

- Selection of morning peak data (07:00 – 09:00)
- Removal of morning peaks with disruptions

- Removal of incomplete AFC transactions
- Inference and increase of occupancy data

Since the aim of our study is to explore how passengers incorporate crowding in their route choice, it is essential that we select a time period in which crowding occurs. Compared to crowding levels reached in metro systems in cities like London, Santiago, Beijing or Tokyo, the level of crowding in The Hague can be considered quite moderate. Outside peak periods, in general no crowding occurs. In peak periods, crowding however does occur on several lines. Since public transport demand in the morning peak in the Netherlands is more concentrated within a relatively small period, compared to a more uniformly distributed demand in the evening peak, we only focus on AFC transactions during the morning peak. This means that only journeys of which the tap-in time lies between 07:00 and 09:00 on working days (Monday – Friday) are considered.

In our study we focus on explaining route choice based on expected attribute values for travel time, waiting time and crowding. Therefore, it is important that only regular, undisrupted periods are incorporated in the dataset. Since disruptions can force passengers to adjust their route choice, this might lead to bias in the analysis. Based on HTM registration data, we removed the AFC data from morning peaks of days where a disruption occurred. Given the possibility of second-order effects, in which a disruption on a certain public transport line might increase occupancies on other parallel lines, we decided to remove a morning peak completely from the data set if a disruption occurred on any line. From the 20 working days remaining in the dataset, we therefore removed the data from 6 working days. For the AFC data of the remaining 14 morning peaks, we removed incomplete transactions.

Transactions can be incomplete due to a system error or due to human error by forgetting to tap-out, in both cases leading to a missing tap-out time and/or location. Although there exist many destination inference algorithms in scientific literature (e.g. the well-known trip chaining algorithm as applied by Trépanier et al. (2007), Zhao et al. (2007) and Wang et al. (2011)), we decided to remove all incomplete transactions. This is because the accuracy of destination inference algorithms for urban tram and bus systems is not 100%. For example, Yap et al. (2017) applied a full validation of the trip-chaining destination inference algorithm, showing that about 65-70% of all destinations are correctly inferred. Since only 1.9% of all AFC transactions are incomplete, there is no necessity to incorporate possibly inaccurate AFC data in this study.

Vehicle occupancies are derived by merging AFC and AVL data. Since both the AFC and AVL data contain the trip number, both data sources can be coupled. This results in the occupancy of each vehicle trip between each pair of stops. These smart card data based occupancies are corrected for the percentage travellers not using a smart card. In the Netherlands, most passengers travel using their smart card. Only passengers who buy a ticket in the vehicle at the driver or vending machine, and passengers who (un)deliberately do not tap in during their trips are not captured. Since our study focuses on experienced crowding levels, these passenger groups are relevant. Based on passenger counts performed by HTM, for each public transport line a correction factor is determined which can be used to increase the smart card based occupancies. This factor varies between 5% and 14% for different lines.

2.2 Transfer inference

For the individual AFC transactions for each trip, it is determined whether an alighting is a transfer or a final destination. To this end, we applied the transfer inference algorithm as described by Yap et al. (2017). This algorithm is an extension on the algorithm described by Gordon et al. (2013), and uses AFC, AVL and inferred occupancy data. Below, this algorithm is shortly discussed. For a more detailed explanation, the reader is referred to Yap et al. (2017). The algorithm contains of a temporal, spatial and binary criterion to categorize an alighting as a transfer:

- Temporal criterion: an alighting is considered a transfer if a passenger boarded the first reasonable vehicle of the next tap-in line passing by the next boarding location. A reasonable vehicle is defined as the first vehicle passing by the next boarding location – given the alighting time of the previous journey leg and required walking time – of which the occupancy does not exceed the norm capacity.
- Spatial criterion: an alighting is considered a transfer if the distance between the alighting location and the next boarding location does not exceed a maximum transfer walking distance of 400 Euclidean meter. An exception to this threshold is made in case passengers use an intermediate public transport service from another operator, of which no AFC data is known.
- Binary criterion: an alighting is not considered a transfer if the next boarding is on the same line as the previous journey stage, since this indicates an activity. An exception is made in case of boarding the first passing vehicle of the same line following the alighted vehicle, since this can indicate (for example) a transfer from a short-service to the long-service vehicle of the same line or a transfer to the same line in case of loops.

In total, the dataset contains 628.839 journeys resulting from 14 working days. Figure 1 (left) shows the resulting distribution of the number of transfers, whereas Figure 1 (right) shows the distribution of journeys over each half-hour period of the morning peak (using the journey tap-in time as classification criterion). As can be seen, almost all journeys consist of 0 or 1 transfer. The busiest part of the morning peak on a network wide level is between 08:00 and 08:30, containing about 31% of all morning peak journeys.

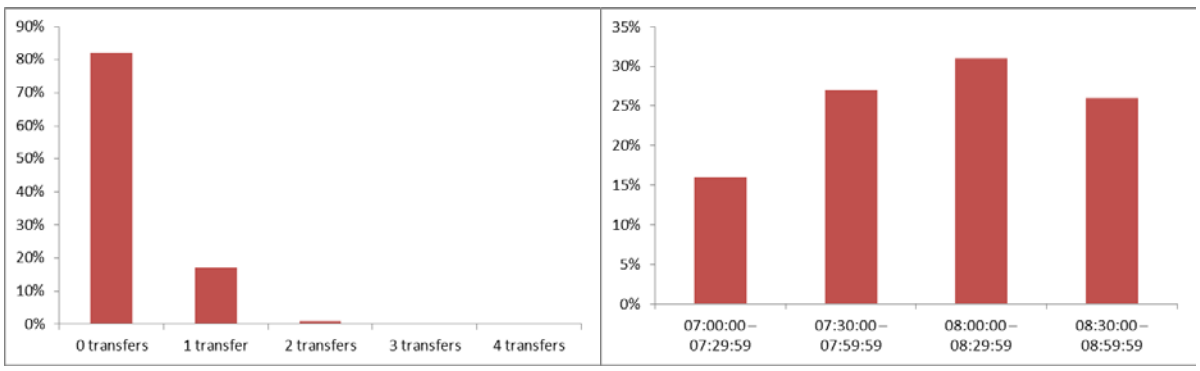

Figure 1. Journey distribution of number of transfers (left) and distribution per half-hour of the morning peak (right)

2.3 Selection of origin-destination pairs

In total, the database consist of 49.231 different chosen routes. This value can be considered as the sum-product of the number of revealed OD pairs and the number of chosen route alternatives for each OD pair. For a given OD pair, a route is considered different in case passengers use a different (combination of) line(s) or use a different transfer location. We developed different criteria to select OD pairs to be used to estimate the discrete choice model with crowding effects:

- Minimum of 2 observed routes per OD pair
- Minimum of 100 observations per OD pair
- Each observed route is chosen by at least 10% of all passengers of that OD pair
- Minimum of 2 observed routes which are physically different from each other
- Existence of a minimum occupancy of 50% on one route
- Attribute variation over all OD pairs

The first criterion is required, since there should be at least an observed choice between two route alternatives. A minimum number of observations for each OD pair is deemed required to incorporate sufficient choices of individual passengers to infer generic coefficients from. Besides, we only incorporate route alternatives which are systematically chosen by a part of the passengers of an OD pair. By setting a minimum value of 10% per chosen route alternative, we ignore route alternatives which are chosen only in a very limited number of cases. Such route alternatives might be chosen in case of disruptions, delays or a-typical passenger behaviour (e.g. passengers travelling for fun). Besides, we checked for each OD pair if at least two route alternatives are physically different from each other. In case of a bundle of different lines sharing the same infrastructure (e.g. for journeys within the city centre), passengers might take the first vehicle passing by suitable for the destination. Given our aim to explain route choice, we consider route alternatives as different from each other, only if these do not share the route for 100%. Another requirement is that there should be at least a certain amount of crowding on (at least) one of the route alternatives of an OD pair. Since we want to estimate crowding valuation, we didn't want to put a crowding constraint *a priori*. However, if no crowding occurs on all route alternatives at all, it is not possible to explain route choice (partly) by crowding. Therefore, we used a very light requirement here. If the expected seat occupancy exceeds 50% during some period of the morning peak on some part of one of the route alternatives, this requirement is fulfilled. A seat occupancy threshold of 50% is used, since passengers start having to sit next to each other from a seat occupancy of 50% or higher. It is expected that negative crowding experiences might start from this value on. At last, we checked over all remaining OD pairs together whether they contain all attributes of the discrete choice model to be estimated. This means that at least some OD pairs should consist of a transfer, tram or bus, in order to be able to estimate the valuation of a transfer or to estimate the in-vehicle time perception in a tram compared to a bus. Applying these criteria results in 58 remaining OD pairs, with in total 17,994 journeys (= 17,994 observations). For all these OD pairs, two route alternatives are observed in the dataset. 16% (9 OD pairs) of these OD pairs consist of a route with a transfer. These 17,994 journeys are made by 7,083 different smart card numbers. Under the assumption that each passenger uses one, unique smart card, this means that there are on average $\approx$2.5 observations per passenger.

2.4 Attributes and attribute levels

In this chapter, we discuss the different attributes and attribute levels for the estimated model.

In-vehicle time
The expected in-vehicle time ivt is determined for each journey leg separately. Based on the AVL data, we calculated the expected in-vehicle time by taking the average realized in-vehicle time over all journeys available in the dataset. Since the scheduled travel times are fixed during the whole morning peak, the average is calculated over the whole morning peak. For each journey leg we introduced a binary variable indicating whether the journey leg is made by tram or bus as mode m.

Waiting time
The expected waiting time wtt expresses the initial waiting time before boarding the first journey leg. Since the AFC system in the Netherlands has only tap-in/tap-out devices within the vehicle, it is not possible to empirically derive the passenger arrival time at the initial boarding stop from the smart card data. In order to quantify the initial waiting time, we assumed a random passenger arrival pattern. Therefore, we set wtt equal to half of the scheduled headway of the boarding line. Since the frequency of each line remains unchanged throughout the morning peak, this value is calculated for the whole morning peak.

Transfer time
The expected transfer time tft expresses the time between the alighting time of the first journey leg and the boarding time of the next journey leg. This means that tft equals the sum of the transfer walking time and transfer waiting time, and equals zero in case the journey consists of one journey leg. The expected value tft for a certain OD pair is calculated by taking the average realized transfer time over all morning peak journeys of that OD pair.

Transfer penalty
tfp is a binary variable which equals 1 in case a journey consists of more than 1 leg, and equals 0 in case a journey consists of 1 leg. This binary variable tfp is used to determine the perceived transfer penalty, which expresses the penalty due to the fact a transfer has to be made, independent from and additional to the perceived transfer time.

Path size
To correct for overlap between route alternatives, we determined the path size factor. The logarithm of the path size factor $lnps$ is used in order to be able to estimate a standard MNL model. The path size factor is computed distance-based: for each route alternative it is determined how many kilometres of the route is shared with the other route alternative of the same OD pair, and how many kilometres are a unique part of only one route alternative.

Crowding: seat occupancy and crowding density
To quantify the valuation of public transport crowding, we use two different attributes: the seat occupancy so and crowding density cd. The attribute values for both attributes are calculated per line, per link (stop-to-stop line segment), per half-hour time period. Given the

non-uniformly distributed demand pattern over the morning peak (Figure 1 right), it is not sufficient to calculate the average values for so and cd over the whole morning peak. Since crowding levels differ over each trip, using a too large time period can result in the use of average crowding levels which do not match with the experienced crowding by passengers during the different trips in the morning peak. On the other hand, passengers will usually not have knowledge of the expected crowding levels for each trip separately. By dividing the morning peak in four periods of 30 minutes, our aim is to find a balance between excluding non-uniformly distributed demand on the one hand, and applying a time period for which passengers can have realistic crowding expectations on the other hand.

The seat occupancy so is calculated using formula (1), and expresses the ratio between the expected passenger load l and the seat capacity c_s. If the expected occupancy exceeds the seat capacity, so remains equal to 1. The expected occupancy is calculated based on the average realized occupancy per stop-stop line segment per half-hour time period. This means that l is calculated as the average realized occupancy over all vehicle trips per 30 minutes. c_s is determined based on data provided by the operator, based on the vehicle type used for each line. To calculate so for each journey leg, the weighted average value is calculated based on the expected seat occupancy and expected travel time ivt_a per link $a \in A$ of the journey leg.

$$so = \min\left(\frac{\sum_a^A \frac{l_a}{c_{s,a}} * ivt_a}{\sum_a^A ivt_a}; 1\right) \tag{1}$$

The crowding density cd is calculated using formula (2), and expresses the expected number of standing passengers per m^2. In line with Wardman & Whelan (2011), we use the crowding density per m^2 instead of the occupancy rate if $l > s_c$, in order to incorporate different vehicle layouts. If the expected passenger load l does not exceed c_s, this value equals zero. This expresses the assumption that passengers will stand only when all seats are occupied. When $l > s_c$, cd is calculated by dividing the number of standing passengers by the total surface available in each vehicle type for standing o. The expected value for cd is calculated by taking the average value per line, per link, per half-hour time period over all realized trips. The expected value per journey leg is computed by using the weighted average over all links.

$$cd = \max\left(\frac{\sum_a^A \frac{l_a - c_{s,a}}{o_a} * ivt_a}{\sum_a^A ivt_a}; 0\right) \tag{2}$$

Table 1 shows the different vehicle types with their corresponding seat capacity c_s and surface available for standing passengers o. Table 2 shows the minimum and maximum value for so and cd per half-hour time period over the total dataset.

Table 1. Seat capacity and standing surface per vehicle type (HTM data)

Vehicle type	Mode	Lines	Seat capacity	Standing surface
GTL-8	Tram	1,6,9,11,12,15,16,17	73	25.1 m^2
Citadis	Light rail	3,4,19	86	32.0 m^2
Avenio	Tram	2	70	33.8 m^2
MAN	Bus	18,21,22,23,24,25,26,28	31	8.9 m^2

Table 2. Min/max seat occupancy and crowding density in dataset

Time period	Seat occupancy		Crowding density	
	Min	*Max*	*Min*	*Max*
07:00 – 07:30	0.10	1.0	0	0.81
07:30 – 08:00	0.08	1.0	0	2.48
08:00 – 08:30	0.13	1.0	0	2.04
08:30 – 09:00	0.11	1.0	0	2.31

2.5 Model formulation

We estimated two models: a model without crowding (model 1) and a model with the incorporation of the two crowding attributes (model 2). Model 2 is an extension of model 1. Formula (3) shows the calculation of the structural utility V. The attributes corresponding to the first journey leg are denoted by index 1; attributes corresponding to the second journey leg are denoted by index 2. As can be seen, generic coefficients are estimated for the initial waiting time and transfer time simultaneously, and for the transfer penalty. Mode-specific coefficients are estimated for in-vehicle time. We experienced with all combinations between estimating only generic coefficients for wtt, ivt, tft, tfp and estimating all mode-specific coefficients for wtt, ivt, tft, tfp. A model using mode-specific in-vehicle time coefficients, and generic waiting+transfer time and transfer penalty coefficients showed to give most reasonable results and the highest value for McFadden's adjusted R^2. From literature, a 'tram bonus' indicating a lower perceived in-vehicle time for tram/rail travelling compared to bus travelled is known (Bunschoten 2013). The selected model allows us to quantify this 'tram bonus' based on RP data as well.

$$V = \alpha_{wtt} * wtt + \alpha_{ivt,m} * ivt_{m,1} + \alpha_{wtt} * tft + \alpha_{tfp} * tfp + \alpha_{ivt,m} * ivt_{m,2} + \alpha_{lnps} * lnps \quad (3)$$

Model 2, being an extension of model 1, estimates the same mode-specific in-vehicle time coefficients and generic waiting+transfer time and transfer penalty coefficients. Formula (4) shows the structural utility calculation when the seat occupancy and crowding density are incorporated. As can be seen, the total in-vehicle time coefficient is now equal to the original in-vehicle time coefficient α_{ivt}, multiplied by a crowding multiplier which is equal to $(1 + (\alpha_{so} * so) + (\alpha_{cd} * cd))$

$$V = \alpha_{wtt} * wtt + (\alpha_{ivt,m} * ivt_{m,1} * (1 + (\alpha_{so} * so_1) + (\alpha_{cd} * cd_1))) + \alpha_{wtt} * tft + \alpha_{tfp} * tfp + (\alpha_{ivt,m} * ivt_{m,2} * (1 + (\alpha_{so} * so_2) + (\alpha_{cd} * cd_2))) + \alpha_{lnps} * lnps \quad (4)$$

Since we incorporated the logarithm of the path size factor in the structural utility component, a correction for overlap between the route alternatives is incorporated in the model. This allows us to estimate standard MNL models as basis. We extended the standard MNL model by incorporating panel effects, since there are multiple route choice observations made by the same smart card number (= the same individual). To correct for the possible correlation between choices made by the same respondent, we estimated a mixed MNL model with panel effects. Formula (5) shows the calculation of the probability for choosing route alternative A. We used Biogeme as software package for performing the maximum likelihood

estimations (Bierlaire 2003). In order to reduce the number of draws, we performed Halton draws from a normal distribution to incorporate the panel structure of the model. In order to determine the number of required Halton draws, we started with an initial number of 5 Halton draws. We doubled the number of Halton draws and checked whether the model outcome can be considered stable. We defined a model outcome as stable, if all estimated coefficients after applying N Halton draws did not significantly differ from the estimated coefficients when $0.5N$ Halton draws were applied. This is the case, when all estimated coefficients do not differ more than twice the standard error of the estimate from the previously estimated coefficients. The model showed to be very stable directly after doubling the number of Halton draws to 10.

$$P_A = \int_{\vartheta_n,\alpha_n} (\prod_{t=1}^{T} (P_{n,A}^{t} | \vartheta_n, \alpha_n) * f(\vartheta_n, \alpha_n)) d\vartheta_n d\alpha_n \quad (5)$$

3. Results and discussion

3.1 Results

Table 3 below shows the number of estimated coefficients, the final log-likelihood, McFadden's adjusted Rho-square and the values of the estimated coefficients with corresponding t-values for model 1 (without crowding) and model 2 (with crowding).

Table 3. Estimation results

	Model 1 (without crowding)	**Model 2 (with crowding)**
Number of observations	17,994	17,994
Number of individuals	7,083	7,083
Number of Halton draws	10	10
Number of estimated coefficients	6	8
Final log-likelihood	-11,404	-11,384
Adjusted Rho-square	0.085	0.087
$\alpha_{ivt,tram}$ (in-vehicle time tram)	-0.158** (-20.6)	-0.151** (-13.7)
$\alpha_{ivt,bus}$ (in-vehicle time bus)	-0.262** (-15.4)	-0.250** (-18.0)
α_{wtt} (waiting+transfer time)	-0.398** (-24.9)	-0.395** (-24.7)
α_{tfp} (transfer penalty)	-0.994** (-9.06)	-1.20** (-10.3)
α_{lnps} (log-path size factor)	2.65** (3.01)	2.37* (2.65)
σ (panel structure)	0.00 (0.00)	0.00 (0.00)
α_{so} (seat occupancy)	-	0.158** (4.97)
α_{cd} (crowding density)	-	0.0611* (2.15)

*t-values in parentheses. * p < 0.05; ** p < 0.01*

Table 4 shows the estimation results for in-vehicle time, waiting+transfer time, transfer penalty and crowding, scaled in which the in-vehicle time coefficient for bus $\alpha_{ivt,bus}$ is set equal to 1.

Table 4. Scaled estimation results

	Model 1 (without crowding)	**Model 2 (with crowding)**
$\alpha_{ivt,tram}$ (in-vehicle time tram)	0.6	0.6
$\alpha_{ivt,bus}$ (in-vehicle time bus)	1.0	1.0
α_{wtt} (waiting+transfer time)	1.5	1.6
α_{tfp} (transfer penalty)	3.8	4.8
α_{so} (seat occupancy)	-	1.16
α_{cd} (crowding density)	-	1.06

From Table 3 we can see that all estimated coefficients are significant, except for the σ coefficient reflecting the panel structure. Apparently, there is no significant correlation between choices made with the same smart card number (made by the same individual). We can also see that the direction of all estimated coefficients is negative, which is plausible. When incorporating crowding in the estimated model, the adjusted Rho-square increases by 2.5% from 0.085 to 0.087. Although the explanatory power of model 2 is only slightly higher than model 1, the LRS-test shows significant results. The LRS-value of 40.6 is larger than the critical χ value of 5.99 (corresponding with 8-6 = 2 degrees of freedom).

From the scaled estimation results for model 2 in Table 4, we can see plausible coefficients. A clear 'tram bonus' can be observed, since 1 minute in-vehicle time by bus is perceived as 0.6 minute in-vehicle time by tram. Earlier research based on SP experiments indicated values between 0.67 and 0.80 (Bunschoten 2013), which means that our research suggest an even heavier 'tram bonus'. One minute waiting time is perceived 1.6 times more negatively, compared to one minute in-vehicle time. This is in line with values found in other studies (e.g. Balcombe et al. 2004). Also the transfer penalty of almost 5 minutes for each (urban) transfer is plausible. In general, we see that RP estimates for the transfer penalty are somewhat lower, compared to SP estimates (e.g. Schakenbos et al. 2016).

Table 5. Crowding multiplier as function of seat occupancy and crowding density

Seat occupancy so (% seats occupied)	**Crowding density (standing pass / m^2)**	**Crowding multiplier**
0	0	1.00
1	0	1.16
1	1	1.22
1	2	1.28
1	3	1.34
1	4	1.40

Table 5 shows the estimated crowding multiplier as function of the seat occupancy and crowding density. As can be seen, the crowding multiplier equals 1.16 when all seats are occupied. In case the occupancy level increases further, the crowding multiplier increases with 0.06 for each increase in the integer number of standing passengers per m^2, additional to the crowding multiplier of 1.16 at seat capacity. In case of on average 4 passengers per m^2, the crowding multiplier thus equals 1.16 + (4 * 0.06) = 1.40. Since we only estimated linear crowding effects, the crowding multiplier in our model increases in linear way from 1.00 to 1.16 when the seat occupancy increases from 0 to 1. However, earlier studies found that the crowding multiplier only starts increasing when the seat capacity is about 0.5-0.8. It is therefore likely that the crowding multiplier starts increasing at a higher seat occupancy than now assumed in our modelling results, and then increases slightly steeper, up to 1.16 when all seats are occupied.

3.2 Discussion

The estimated coefficients show plausible results which are in line with values found in earlier studies. Also when crowding is incorporated in the model, results seem plausible. The crowding multiplier found in our study is lower than values found in SP experiments, for example as reported by MVA Consultancy in Wardman & Whelan (2011). In general, there is a tendency that values found using SP experiments are higher than values found in RP studies. In stated situations, respondents are apparently more inclined to incorporate crowding in their route choice, compared to their realized behaviour in practice. When we compare our study results with the RP results found for the Hong Kong MTR metro by Hörcher et al. (2017), we see that the multipliers found in our study are somewhat lower. However, differences between our study and this RP based study are clearly less, compared to differences found between our study and previous SP based studies.

In this study, we used the average realized occupancies per line, per link, per half-hour time period as expected crowding level. We might hypothesize that the experienced level of crowding differs somewhat from the expected crowding level we used in this study, due to irregularity and unreliability. Since unreliability and bunching causes more passengers to experience more crowding, while less passengers experience less crowding than average, there might be some discrepancy between perceived and assumed expected crowding level here. We also assume an equal distribution of passengers throughout the vehicle. In reality we see that not all locations in the vehicle are equally suitable for standing passengers, for example because of the lack of holding possibilities in some parts of the vehicle. Therefore, it might be the case that crowding levels in some parts of the vehicle as substantially higher than average. This can also lead to a discrepancy between perceived and in our study assumed expected crowding levels.

Besides, in our study we did not incorporate any segmentation. It can be expected that especially frequent travellers have good expectations regarding crowding levels based on their prior travel experiences, whereas infrequent travellers with limited prior experiences might hardly incorporate crowding in their route choice. In case a distinction between frequent and infrequent travellers would be made in the estimations, it might be possible that the crowding multiplier during the morning peak (with usually a relatively high share of frequent, experienced travellers) increases.

At last, an important limitation for this study is that no information is available regarding the realized passenger arrival time at the stop. Our study assumes that passenger route choice is fully based on expected crowding levels. Due to the lack of this information, we cannot determine whether a passenger boarded the first vehicle passing the stop, or deliberately skipped a crowding vehicle for a less crowded alternative. Information about this would enable us to investigate to which extent the real-time crowding level of the arriving vehicles affects route choice, compared to expected crowding levels based on prior experiences.

4. Conclusions

Crowding in public transport can be of major influence on passengers' travel experience and therefore affect route and mode choice. Because of the expected increasing concentration of activities within urban agglomerations in many countries worldwide, crowding is expected to become an even more dominant factor in urban public transportation in the future. Therefore, it is important to understand how crowding in urban public transport is perceived by passengers. Over the years, there are several studies performed aiming to better understand the valuation of crowding in public transport. The majority of these studies use stated preference (SP) experiments to estimate crowding valuation. It is however known that there can be a discrepancy between stated choices of respondents in SP experiments, compared to revealed choice behaviour in reality. However, the availability of automatic fare collection (AFC), automatic vehicle location (AVL) and automatic passenger count (APC) systems in public transport the last years enables the estimation of crowding valuation based on large scale revealed preference data. The availability of individual smart card transactions allows us to gain insights in revealed trade-offs between (for example) travel time, transfers, waiting time and crowding in public transport route choice.

Based on the estimated discrete choice model, we conclude that crowding plays a significant role in passengers' route choice in public transport. The crowding multiplier of in-vehicle time equals 1.16 when all seats are occupied. In case the occupancy level increases further, the crowding multiplier increases with 0.06 for each increase in the integer number of standing passengers per m^2, additional to the crowding multiplier of 1.16 at seat capacity. In case of on average 4 passengers per m^2, the crowding multiplier thus equals 1.16 + (4 * 0.06) = 1.40.

The estimated models in our study show plausible results. Besides, our study is the first in which crowding valuation for urban tram and bus travelling is determined fully based on revealed preference data. The insights gained from our study can support the decision-making process, by quantifying benefits of measures aiming to reduce crowding levels. The estimation of segment models, distinguishing between (in)frequent travellers or between different travel purposes, is recommended for further research.

Acknowledgements

This research was performed as part of the TRANS-FORM (Smart transfers through unravelling urban form and travel flow dynamics) project funded by NOW grant agreement 438.15.404/298 as part of JPI Urban Europe ERA-NET CoFound Smart Cities and Communities initiative. The authors thank HTM, the urban public transport operator of The Hague, the Netherlands, for their valuable cooperation and data provision.

References

Balcombe, R., Mackett, R., Paulley, N., Preston, J., Shires, J., Titheridge, H., Wardman, M. & White, P. (2004). *The demand for public transport: a practical guide*. TRL Report TRL 593, Crowthorne, UK.

Bierlaire, M. (2003). BIOGEME: A free package for the estimation of discrete choice models. *Proceedings of the 3rd Swiss Transportation Research Conference,* Ascona, Switzerland.

Bunschoten, T., Molin, E.J.E. & Van Nes, R. (2013). Tram or bus; does the tram bonus exist? *Proceedings of the European Transport Conference*, Frankfurt, Germany, AET.

Cats, O., West, J. & Eliasson, J. (2016). A dynamic stochastic model for evaluating congestion and crowding effects in transit systems. *Transportation Research Part B,* 89, 43–57.

Gordon, J.B., Koutsopoulos, H.N., Wilson, N.H.M. & Attanucci, J.P. (2013). Automated inference of linked transit journeys in London using fare-transaction and vehicle location data. *Transportation Research Record: Journal of the Transportation Research Board,* 2343, 17-24.

Hamdouch, Y., Ho, H.W., Sumalee, A. & Wang, G. (2011). Schedule-based transit assignment model with vehicle capacity and seat availability. *Transportation Research Part B,* 45, 1805–1830.

Haywood, L. & Koning, M. (2015). The distribution of crowding costs in public transport: new evidence from Paris. *Transportation Research Part A,* 77, 182–201.

Hörcher, D., Graham, D.J., Anderson, R.J. (2017). Crowding cost estimation with large scale smart card and vehicle location data. *Transportation Research Part B,* 95, 105-125.

Kroes, E., Kouwenhoven, M., Debrincat, L. & Pauget, N. (2014). On the value of crowding in public transport for Ile-de-France. *Transport Research Arena,* Paris, France.

Li, Z. & Hensher, D. (2011). Crowding and public transport: A review of willingness to pay evidence and its relevance in project appraisal. *Transport Policy,* 18, 880-887.

Munizaga, M.A. & Palma, C. (2012). Estimation of a disaggregate multimodal public transport origin-destination matrix from passive smartcard data from Santiago, Chile. *Transportation Research Part C,* 24, 9-18.

Nuzzolo, A., Crisalli, U. & Rosati, L. (2012). A schedule-based assignment model with explicit capacity constraints for congested transit networks. *Transportation Research Part C,* 20, 16–33.

Pel, A., Bel, N., Pieters, M. (2014). Including passengers' response to crowding in the Dutch national train passenger assignment model. *Transportation Research Part A,* 66, 111-126.

Prud'homme, R., Koning, M., Lenormand, L. & Fehr, A. (2012). Public transport congestion costs: the case of the Paris subway. *Transport Policy*, 21, 101–109.

Schakenbos, R., La Paix, L., Nijenstein, S. & Geurs, K. (2016). Valuation of a transfer in a multimodal public transport trip. *Transport Policy,* 46, 72-81.

Schmöcker, J.-D., Fonzone, A., Shimamoto, H., Kurauchi, F. & Bell, M.G.H. (2011). Frequency-based transit assignment considering seat capacities. *Transportation Research part B,* 45, 392–408.

Trépanier, M., Tranchant, N. & Chapleau, R. (2007). Individual trip destination estimation in a transit smart card automated fare collection system. *Journal of Intelligent Transportation Systems,* 11, 1-14.

Van Oort, N., Brands, T., de Romph, E. & Yap, M.D. (2016). *Ridership Evaluation and Prediction in Public Transport by Processing Smart Card Data: A Dutch Approach and Example.* Chapter 11*, Public Transport Planning with Smart Card Data,* eds. Kurauchi F., Schmöcker, J.D., CRC Press.

Wang, W., Attanucci, J.P. & Wilson, N.H.M. (2011). Bus passenger origin-destination estimation and related analyses using automated data collection systems. *Journal of Public Transportation,* 14, 131-150.

Wardman, M. & Whelan, G. (2011). Twenty years of rail crowding valuation studies: evidence and lessons from British experience. *Transport Reviews,* 31, 379-398.

Yap, M.D., Cats, O., Van Oort, N. & Hoogendoorn, S.P. (2017). Robust transfer inference: a transfer inference algorithm for public transport journeys during disruptions. *Prepared for the 20th EURO Working Group on Transportation Meeting 2017, Budapest, Hungary.*

Zhao, J., Rahbee, A. & Wilson, N.H.M. (2007). Estimating a rail passenger trip origin-destination matrix using automatic data collection systems. *Computer-Aided Civil and Infrastructure Engineering,* 24*,* 376-387.